Visualization Tool
for
Flow Cytometry Data Standards
Project

Evgeny Maksakov
maksakov@cs.ubc.ca

CS533C
Department of Computer Science, UBC
in collaboration with
Terry Fox Laboratory, BC Cancer Agency
(Prof. Ryan Brinkman & Dr. Josef Spidlen)
Today

• Flow Cytometry (reminder)
  – Dataset description
• Goals
• Previous work
• FlowCytoVis prototype in details
• Data analysis comparison
  – FlowJo vs FlowCytoVis prototype
• Demo!
• Conclusions and future work
Flow Cytometry (FCM)

Cell Measure

FACSCalibur Optical Layout
Dataset Properties

Typically for research at the TFL:
• 100,000+ events
• 5-10 dimensions

Capability:
• 1,000,000 events (cells going through the laser beam) per dataset
• Up to 20 dimensions

Today demo datasets:
• 20,000 events
• 5 dimensions
Dimensions

SSC (Granularity)  
PI dye intensity (measures viability)

Laser  
FSC (Size)

Green Fluorescent Protein intensity (measures gene expression)

16 fluorescence intensities of fluorochromes (used as markers)

Pictures are taken from http://www.upenn.edu/pennnews/photos/, http://www.bdbiosciences.com/image_library/ and flow cytometry manual
Aimed Goals

User requirements (based on user studies):
1. See all dimensions at once
2. Improve analysis sequence
3. Leave scatterplots and histograms
4. Gating/Filtering feature
5. Provide better usability than commercial FlowJo

By means of:
1. Using Parallel Coordinates with Gating/Filtering
2. Implementing data clustering throughout dimensions
3. Include scatterplots and histograms in the interface
4. Make effective, convenient and interactive interface
3D Parallel Coordinate System for FCM
Marc Streit at al. (2006)

---

*Picture from Marc Streit at al. (2006)*
3D Parallel Coordinate Problems

- Does not provide any new information about dataset
- Introduces visual occlusions
- Necessity to rotate to see all data
FlowCytosis

- FCM Data
- Interchangeable dimensions
- Clusters
- Gates/Filters
- Collapsing axes captions
- Histograms and Scatterplot Buttons
- Dataset tabs

Statistics
- Events count: 200000
- Shown (% of total): 35.29
- Dropped (% of shown): 26.4
- Clusters found: 45
Aimed Goals

User requirements (based on user studies):

1. See all dimensions at once
2. Improve analysis sequence
3. Leave scatterplots and histograms
4. Gating/Filtering feature
5. Provide better usability than commercial FlowJo

By means of:

1. Using Parallel Coordinates with Gating/Filtering
2. Implementing data clustering throughout dimensions
3. Include scatterplots and histograms in the interface
4. Make effective, convenient and interactive interface
Data Analysis Process (FlowJo)

Negative control
(each scatterplot is a new window)

Event Count: 28988

Event Count: 18229

Event Count: 17755

Event Count is a total number of cells passed through the laser beam

Important note: sequence of actions is the same all the time for negative control!
Data Analysis Process (FlowCytoVis)

Negative control
(everything happens in one window)
Data Analysis Process (FlowJo)

Looking for result

Event Count: 16061

Non-marked cells

Marked cells (result)

Important note: Same gates as in neg. control apply automatically on the positive set!
Data Analysis Process (FlowCytoVis)

Looking for result

Important note: Gates apply automatically on the positive set here too!
Aimed Goals

User requirements (based on user studies):

☑ 1. See all dimensions at once
☑ 2. Improve analysis sequence
☑ 3. Leave scatterplots and histograms
☑ 4. Gating/Filtering feature

? 5. Provide better usability than commercial FlowJo

By means of:

☑ 1. Using Parallel Coordinates with Gating/Filtering
☑ 2. Implementing data clustering throughout dimensions
☑ 3. Include scatterplots and histograms in the interface
☑ 4. Make effective, convenient and interactive interface
Demo

Implementation details:

- Java2D + Swing
- CFCS library for reading .fcs (FCM datasets) format
Strengths and Weaknesses of the FlowCytoVis

+ Can provide insights into the data
+ Convenient (less clicks to get the same result)
+ Interactive
+ Allows intuitive multidimensional filtering
+ Visually appealing

- Slow picture rendering relatively to Scatterplots
- At the moment does not provide full functionality that FlowJo provides.
Conclusions

• The FlowCytoVis proved to be a relevant solution for the Flow Cytometry data visualization and was accepted with enthusiasm

• Parallel Coordinates (PC) view is a nice addition to canonical Scatter Plots for Flow Cytometry

• Clustering works very well together with PC and can save some rendering time

• Clustering needs refinement and improvement

• Improving speed is vital for PC
Future Work

• Implement all the functionality still missing

• Integrate existing clustering made for the Flow Cytometry Data Standards Project into the FlowCytoVis

• Improve rendering speed for parallel coordinates
Acknowledgements

• Dr. Tamara Munzner
• Dr. Ryan Brinkman
• Dr. Josef Spidlen
• Dr. Louie van de Lagemaat
• Irina Maksakova
• Other TFL Members
Questions...